

## Article

# The Value of Twins for Health and Medical Research: A Third of a Century of Progress

Jeffrey M. Craig<sup>1,2</sup>, Lucas Calais-Ferreira<sup>3,4</sup>, Mark P. Umstad<sup>5,6</sup> and Dedra Buchwald<sup>7</sup>

<sup>1</sup>Centre for Molecular and Medical Research, School of Medicine, Faculty of Health, Deakin University, Geelong, VIC, Australia, <sup>2</sup>Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, VIC, Australia, <sup>3</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia, <sup>4</sup>CAPEs Foundation, Ministry of Education, Brasilia, Brazil, <sup>5</sup>Department of Maternal–Fetal Medicine, Royal Women's Hospital, Parkville, VIC, Australia, <sup>6</sup>Department of Obstetrics and Gynaecology, University of Melbourne, Melbourne, VIC, Australia and <sup>7</sup>Elson S. Floyd College of Medicine, Institute for Research and Education to Advance Community Health, Washington State University, Spokane and Seattle, WA, USA

## Abstract

In 1984, Hrubec and Robinette published what was arguably the first review of the role of twins in medical research. The authors acknowledged a growing distinction between two categories of twin studies: those aimed at assessing genetic contributions to disease and those aimed at assessing environmental contributions while controlling for genetic variation. They concluded with a brief section on recently founded twin registries that had begun to provide unprecedented access to twins for medical research. Here we offer an overview of the twin research that, in our estimation, best represents the field has progress since 1984. We start by summarizing what we know about twinning. We then focus on the value of twin study designs to differentiate between genetic and environmental influences on health and on emerging applications of twins in multiple areas of medical research. We finish by describing how twin registries and networks are accelerating twin research worldwide.

**Keywords:** Twins; medical research; statistical models; genetics; environment; stem cells

(Received 19 November 2019; accepted 26 November 2019; First Published online 27 January 2020)

## Twin Types

In the context of twin research, zygosity refers to the genetic similarity between members of a twin pair. Monozygotic (MZ) twins arise from a single zygote, which splits into two after a number of cell divisions (Hall, 2003); these twins are often described as 'identical'. Dizygotic (DZ) twins arise from two separate zygotes and are often described as 'fraternal' or 'non-identical' (Figure 1). DZ twins are no more or less genetically similar than any two siblings from the same parents; on average, they share 50% of their genetic variation. Approximately half of DZ twin pairs are the same sex and half are opposite sexes. Like singletons, each member of a DZ pair develops its own chorion (thick outer sac continuous with the placenta, Figure 1). Members of 'dichorionic' twin pairs develop in two placentas, while 'monochorionic' twins share the same placenta (Hall, 2003; McNamara et al., 2016). All DZ twins are dichorionic, but approximately two-fifths of MZ twins are dichorionic and three-fifths are monochorionic. About 1% of MZ twins are both monochorionic and monoamniotic (sharing a single amniotic sac) and are known popularly as 'MoMo' twins. Certain rare types of twins and twin-related phenomena, as well as the origin of twins, are discussed in detail elsewhere (Hall, 2003; McNamara et al., 2016).

## The Value and Variety of Twin Study Designs

With rare exceptions, findings from twin studies are applicable to singletons. For example, twins and singletons share a similar incidence of chronic diseases (Martin et al., 1997) and neurodevelopmental disorders (Lorenz, 2012). These similarities underscore the value of twin research for the population as a whole. However, monochorionic twins have a higher rate of morbidity and mortality than singletons, whereas twins of all types tend to have shorter gestation and lower birth weight than singletons. Although these factors have little effect on chronic disease incidence, studies of neurodevelopmental disorders are advised to adjust for them (Lorenz, 2012). To date, cerebral palsy is the only condition with a higher incidence in twins after adjusting for gestational age and birth weight (Bonellie et al., 2005), prompting caution when findings on this condition in twins are extrapolated to singletons.

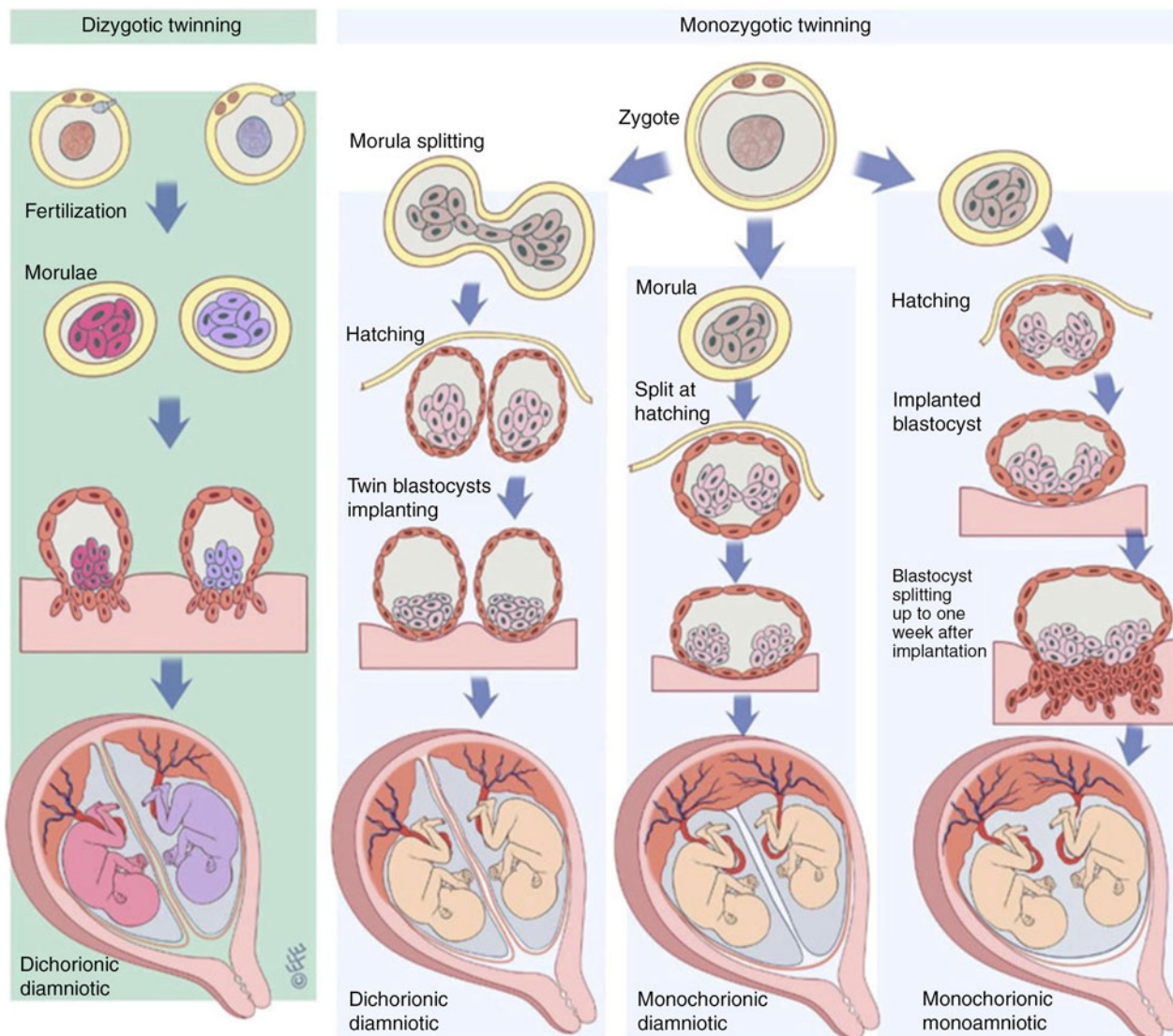
## Classical Twin Models

Various analytic models have capitalized on characteristics unique to twins (reviewed in Boomsma et al. (2002)). Early twin studies compared phenotypic similarity within MZ twin pairs to that within DZ pairs to establish the proportion of variance in a given trait that is due to genetic variation within a population (also known as heritability) and the proportion due to environmental variation. Such heritability studies are often called 'classical twin models'. In 1993, using a classical twin model with both MZ and DZ pairs, Berkovic et al. (1993) confirmed an earlier finding

**Author for correspondence:** Jeffrey M. Craig, Email [Jeffrey.craig@deakin.edu.au](mailto:Jeffrey.craig@deakin.edu.au)

**Cite this article:** Craig JM, Calais-Ferreira L, Umstad MP, and Buchwald D. (2020) The Value of Twins for Health and Medical Research: A Third of a Century of Progress. *Twin Research and Human Genetics* 23: 8–15, <https://doi.org/10.1017/thg.2020.4>

© The Author(s) 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Fig. 1.** The formation of dizygotic and monozygotic twins. Dizygotic twins are the product of two fertilization events resulting in dichorionic diamniotic twins with each twin developing to become a genetically distinct individual. Monozygotic twins result from post-zygotic splitting of the product of a single fertilization event. Dichorionic diamniotic twins result from splitting up to the morula stage, monochorionic diamniotic twins from splitting at the blastocyst hatching stage and monochorionic monoamniotic from splitting at the implantation stage. Reproduced with permission from McNamara et al. (2016).

that epilepsy is highly heritable and challenged the causal role of perinatal factors in this illness. They also found that twins and singletons have a similar risk of epileptic seizures, a result with important implications for medical practice. A few years later, a study using a similar model calculated a high heritability for autism spectrum disorders (Bailey et al., 1995) which had previously been ascribed to lack of maternal warmth.

Classical twin models can also estimate, for a given phenotype, the proportion of variance due to genetics (the sequence of DNA itself), shared environment (all exposures that twins share) and nonshared environment (exposures and stochastic effects that differ within pairs). Such models were in their infancy 33 years ago. They use statistical techniques in which phenotypic variance is partitioned into variance due to additive genetic effects (abbreviated A), common or shared environment (C) and nonshared or unique environment (E) (Boomsma et al., 2002). These three types of variance are usually expressed as ratios that sum to one.

'Additive' implies that the number of genetic variants correlates with the severity of the phenotype in a linear manner. Additive factors, along with other genetic effects such as dominance, are used in estimating heritability. Heritability is a proportion rather than an absolute number, while genotype is largely fixed at conception. Therefore, the estimate of heritability is inversely proportional to the total variation in phenotype that can be increased by geographic and age-related factors. For example, height is less heritable in developing countries with higher levels of environmental adversity than in developed countries with less adversity (Silventoinen et al., 2003), while the heritability of body mass index (BMI) increases from birth to childhood and then declines in adulthood (Elks et al., 2012). These and other issues have led many researchers to advise caution when interpreting heritability estimates (Hopper, 1993; Turkheimer, 2011). Details of other types of statistical methods applied to the analysis of variance components in twin studies are reported elsewhere (Boomsma et al., 2002; Hopper, 1993).

## Twin Differences Models

'Twin differences' models, which focus on comparisons within MZ pairs, were developed in the early 1940s but have not been used as extensively as classical twin models. Within-pair analysis of MZ twins controls for paternity, maternal factors during pregnancy, gestational age, location and season of birth, post-natal familial factors, age, ethnicity, sex and genetics (Hall, 2003; McNamara et al., 2016). Within-pair analysis of DZ twin pairs controls for all these factors with the exception of genetics, for which only 50% of variation is controlled, and sex that is controlled only in same-sex pairs.

Although within-DZ pair designs have been used less often than MZ designs, one approach is worth highlighting. Within-pair analysis of opposite-sex DZ pairs can address sex differences in the risk, severity, outcomes and response to intervention of a specific disease or disorder. For example, a study of opposite-sex DZ pairs concordant for autism spectrum disorder (Robinson et al., 2013) found that boys had more severe symptoms than their twin sisters. These results corroborated a previous hypothesis that female sex could protect against autism. Other notable applications of this approach include findings that males are more vulnerable to delinquency than females (Newsome et al., 2016) and that males and females differ on some causes of depression (Kendler & Gardner, 2014).

The co-twin control model is a type of twin differences model that focuses exclusively on MZ pairs who are discordant either for a disease or for an exposure, such as an environmental factor. This model has recently been used to investigate environmental contributions to the associations between BMI and cardiovascular disease (Song et al., 2015), between height and asthma (Protudjer et al., 2015) and between pain risk factors and low back pain (Oliveira et al., 2015). One notable study with this design demonstrated that smoking was associated with a reduction of 5–10% in bone density, thereby conferring a high risk of fracture (Hopper & Seeman, 1994).

## Clinical Trials

An especially efficient version of the co-twin control model has been used in clinical trials, usually those aimed at establishing the effectiveness of a medical intervention. This approach can use either MZ pairs exclusively or MZ and same-sex DZ pairs in combination. One member of each twin pair is randomized to the treatment group and the other to the control group, and then the association between exposure and outcome is analyzed from the perspective of within-pair differences. This model can achieve up to seven times more statistical power to detect an intervention effect than can traditional randomized controlled trials (Carr et al., 1981)

Unfortunately, twins remain underrepresented in clinical trials (Yelland et al., 2015). Even when they are included, analytical methods are often inadequately described or do not account for twin clustering, which is a function of the common environment and genetic background shared by each twin pair (Hibbs et al., 2010). One notable exception was a meta-analysis of 34 randomized controlled trials with an aggregated sample of 14,106 singletons and 2578 twins (both MZ and DZ; Saccone et al., 2015) This analysis found insufficient evidence to support the routine use of Omega-3 supplements during pregnancy to prevent preterm birth and other complications, such as preeclampsia. Another example is an ongoing trial to test an intervention to improve sleep quality in patients with low back pain (Pinheiro et al., 2018). A further example is a clinical trial that found that dietary fat intake is the main influencer of fat taste sensitivity, which is

a marker of satiety and a contributory factor for obesity (Costanzo et al., 2018).

## Gene–Environment Interactions

Gene–environment *correlation* occurs when exposure to a given environment is associated with a specific genotype. Gene–environment *interaction* occurs when people with different genotypes respond to the same environment in different ways. Studies investigating such phenomena are essential to understanding the interplay between genes and environment, and they have great potential to inform the personal tailoring of medical care. A 2015 study of MZ and DZ twins found that genetic effects on musical accomplishment were most pronounced among musicians who practiced frequently (Hambrick & Tucker-Drob, 2015). Another study found that a genetic disposition to depression in MZ and DZ co-twins significantly interacted with environmental triggers, such as stressful life events, to precipitate depressive episodes (Kendler et al., 1995). Another, using an unspecified number of MZ and DZ twin pairs, found that gene–environment interactions can influence levels of gene expression (Buil et al., 2015).

## Cause Versus Association

Some twin models are well suited to studies that seek to differentiate between cause and association in human health (reviewed in Hrubec & Robinette, 1984; and van Dongen et al., 2012). Their sensitivity stems from the unique capacity of twin designs, especially the MZ co-twin design, to control for variables that often confound population studies. For example, a study with 1196 MZ and 1352 DZ twin pairs showed that even though regular exercise is associated with reductions in symptoms of anxiety and depression in the general population, the association is not due to the effects of exercise (De Moor et al., 2008). This conclusion was based on findings that MZ twin pairs discordant for physical activity reported similar levels of anxiety and depression. A study with a similar design in 157 MZ twin pairs found that later bedtimes and shorter sleep duration predicted subsequent development of depression, anxiety and self-injury (Matamura et al., 2014). A longitudinal study of 2464 elderly Danish twins noted associations among smoking, BMI and longevity, but concluded that the links between smoking and mortality and between BMI and mortality were more likely to result from environmental effects than from genetics (Herskind et al., 1996). More recently, a study of 503 MZ twin pairs found a causal relationship between measures of stress and tension and subsequent depression and anxiety (Davey et al., 2016). Therefore, interventions that successfully minimize stress and tension should also have positive effects on depression and anxiety.

## Emerging Uses of Twins in Medical Research

### Epigenetics and Twins

Epigenetics is the study of changes in an organism over time caused by the addition and removal of small molecules in DNA and DNA-packaging proteins, without any changes in DNA sequence. Well-studied examples of epigenetic processes include the covalent addition of a methyl group (CH<sub>3</sub>) to the cytosine nucleotide of a cytosine–guanine sequence (CpG site) and the covalent addition of an acetyl group (C<sub>2</sub>H<sub>3</sub>O) to DNA-packaging histone proteins. Epigenetic changes are intrinsic to mammalian development and are often susceptible to influence by environmental conditions that can be internal (e.g., fetal nutrient availability) or external (e.g., socioeconomic status; Chiarella et al., 2015). Studies involving



genome-wide analysis of DNA methylation have begun to identify causal mechanisms and diagnostic biomarkers for complex human diseases (Mikeska & Craig, 2014).

Twin studies in particular can yield valuable insights in epigenetics and its application to the epigenome — that is, the sum total of epigenetic marks in any given tissue. This research has been reviewed elsewhere (Bell & Saffery, 2012; Chiarella et al., 2015; Martin et al., 1997; van Dongen et al., 2012), so we limit our focus to epigenome-wide studies involving twins, especially studies of the etiology of human disease and the variance components of DNA methylation itself. We discuss the role of epigenetics in the developmental origins of disease in the next section.

Epigenome-wide association studies analyze the associations between DNA methylation and a wide variety of environments, phenotypes and diseases (Bell & Saffery, 2012; Martin et al., 1997; van Dongen et al., 2012). Most have used microarrays (DNA sequences laid out on glass slides) that enable quantification of DNA methylation in regions of the genome that regulate gene activity. Given the possibility of environment- and phenotype-discordant co-twin designs, epigenome-wide association studies with MZ twins offer greater power than studies with singletons, because MZ twins present essentially no variation in the DNA sequence itself. Such studies are recommended as a first step in the epigenetic analysis of human disease phenotypes (Rakyan et al., 2011), and notable recent examples include epigenome-wide association studies of twins discordant for cerebral palsy (Mohandas et al., 2018), multiple sclerosis (Souren et al., 2019) and BMI (Li et al., 2019).

However, epigenome-wide association studies with exposure-discordant pairs are uncommon, and to our knowledge, the only study of this kind performed to date involved 20 MZ twin pairs discordant for smoking (Allione et al., 2015). Even with this limited sample, the study identified eight loci previously associated with smoking in singletons.

The DNA methylation level at each CpG site in the genome in any given tissue can be treated as a separate phenotype and analyzed accordingly by statistical modeling, with the caveat that levels of DNA methylation within a region of hundreds of base pairs are often highly correlated. Studies using this approach have revealed a wide range in the influence of genetic, shared and nonshared environments on variation in DNA methylation. Using microarrays, researchers have found that the average heritability of DNA methylation across the genome is relatively low (McRae et al., 2014; van Dongen et al., 2016). A recent study of DNA methylation at more than 410,000 CpG sites in peripheral blood leukocytes of 769 MZ and 424 DZ twin pairs found a heritability of .19, a nonshared environmental component of .81 and negligible evidence for an effect of shared environment (van Dongen et al., 2016). Another study using the same microarray technology and tissue with 614 participants, including 69 MZ and 11 DZ twin pairs, found that highly heritable CpG sites were influenced by genetic variants within hundreds of base pairs (McRae et al., 2014). A much smaller study using more than 19,000 CpG sites in multiple tissues from 22 MZ and 12 DZ newborn twin pairs also found that the nonshared environment was by far the largest contributor to DNA methylation (Gordon et al., 2012). The key role played by the nonshared environment suggests that some combination of stochastic factors, including random biological variation or 'noise' (Little et al., 2013) and twin-specific factors, is the most common influence on the epigenome. The likelihood that most of this variation happens in utero is reviewed in the next section.

## The Developmental Origins of Health and Disease

Accumulated evidence from human and animal studies has shown beyond reasonable doubt that most of the environmental components of chronic illness, including cardiometabolic and psychiatric diseases and some cancers, originate early in life (Gluckman et al., 2010). It is also understood that humans are affected by the environment throughout the lifespan, although this influence diminishes over time. The field of developmental origins of health and disease, as it is widely known, arguably originated with the discovery of the link between low birth weight (a proxy for intrauterine growth) and elevated risk of cardiometabolic disease (Barker & Osmond, 1988). This field has since expanded to include such concepts as developmental mismatch, in which the prenatal maternal environment to which a fetus is exposed differs markedly from the postnatal environment (Gluckman et al., 2010).

Twins have played a prominent role in this research. At a fundamental level, twin studies enable the partitioning of variance in phenotypic outcomes (such as low birth weight) into genetic, shared and nonshared environments. Only twin research can differentiate between the latter two. Except in rare circumstances, all twin pairs share the same mother and father. In addition, the shared environment to which twins are exposed in utero includes aspects of maternal diet and lifestyle, since soluble factors in the mother's body — including nutrients, oxygen, inflammatory factors, some infectious agents and hormones — are passed to both twins *via* the placenta and umbilical cords. All twins also share the same season of birth, and most share the same early postnatal family environment. Studies of environmental risk factors involving siblings or unrelated singletons are unable to adequately control for these shared factors.

Nonetheless, despite the important contribution of shared factors to chronic disease risk, nonshared factors are likely to be more influential, as they are in epigenetics. A recent meta-analysis of data on 28 chronic diseases from more than 150,000 European twins found that the sum of shared genetics and shared environment accounted on average for only 18.5% of variance in disease phenotype (Rappaport, 2016). A recent in-depth study of more than 200 immune parameters in 201 healthy twins found that 58% of these parameters were almost completely determined by nongenetic factors and that nonshared environment dominated (Brodin et al., 2015). One seminal study (Plomin & Daniels, 1987) from the 1980s highlighted the high rate of discordance for behavioral phenotypes within MZ pairs, concluding that many factors must be capable of producing health differences between identical twins. Research over the ensuing decades has addressed the intrauterine environment as a critical source of influences on the development of individual members of a twin pair (Martin et al., 1997; Stromswold, 2006).

In almost all twins, the placenta and umbilical cords can be classified as contributors to the nonshared environment because of their physical and biological variability within pairs. Such variation can arise from differences in the uterine implantation site; in the length, width, torsion, knotting and number of vessels of the umbilical cord; and in the location where the umbilical cord is inserted. All these factors have the potential to affect the growth rate and the rate of transplacental transport of oxygen, nutrients and teratogens in individual twins. In addition, any twin pair can be discordant for prenatal (Dickinson et al., 2006; Pimentel et al., 2012) or perinatal (Jamieson et al., 2007) infection, whether viral (Dickinson et al., 2006) or bacterial (Pimentel et al., 2012) or concordant for infection but discordant for severity of symptoms

(Schiesser et al., 2009). Inflammation of the umbilical cord (funisitis) and the placenta (chorioamnionitis) can be confined to the placenta of a single twin (Dickinson et al., 2006). If infectious agents enter the embryo *via* the placenta and cord, this process might be a mechanism for discordant placental inflammation, which would lead to discordant prenatal infection. Another potential source of discordance could be the ability of each placenta to induce a separate maternal immune response (Yusuf & Kliman, 2008). Once an infection appears in the amniotic sac of one twin, layers of amnion — and in dichorionic twins, of chorion, in addition to the co-twin's placenta — could protect the co-twin from its spread.

Hrubec and Robinette (1984) reported that MZ co-twin studies had already revealed associations linking low birth weight with cerebral palsy as well as schizophrenia. Since then, studies using this design have found associations linking higher birth weight with cognition in very early childhood (Halling et al., 2015) and with higher BMI in adolescence (Yokoyama et al., 2016) and adulthood (Pietilainen et al., 2002), as well as associations linking lower birth weight with higher blood pressure in childhood (Dwyer et al., 1999), with early onset of puberty (Schulte et al., 2016) and with risk of type 2 diabetes in adulthood (Poulsen et al., 1997). Dwyer et al. (1999) concluded that the critical step in the causal pathway from low birth weight to hypertension must involve mechanisms in the placenta, cord and fetus, with maternal exposures operating only as antecedents or primers. Collectively, these data add to the evidence that prenatal factors have long-lasting effects on health. In addition, birth weight is correlated with placenta weight, with a higher correlation within pairs than between pairs (Gielen et al., 2007). This again suggests that nonshared factors have more bearing on later health outcomes than do shared factors. In a within-pair model, higher birth weight has also been associated with a more central umbilical cord insertion (Gielen et al., 2006), which in turn has been linked with more favorable long-term health outcomes (Antoniou et al., 2011).

Epigenomic studies have also been used to study the early-life origins of disease in twins. One set of studies found that at birth, the epigenetic mark of DNA methylation exhibits a range of within-pair variation throughout the genome (Gordon et al., 2012). Results also showed that a small number of MZ twin pairs are more epigenetically discordant even than some unrelated individuals, emphasizing the formative nature of time spent in the womb. In addition, these studies have found that within-pair differences in DNA methylation and expression of genes involved in lipid metabolism, immune function and cardiometabolic function are associated with birth weight discordance. This finding is consistent with the previously established role of DNA methylation in mediating the link between low birth weight and chronic disease risk (Gluckman, 2012).

## Microbiota

Recent research using DNA sequencing has revealed the existence of a large number of previously unrecognized species of microbiota that cannot be grown in culture. This work has identified roles for microbiota in regulating metabolism, immunity and brain function (Spector, 2015). Central to this developing field is the question of whether genetic or environmental variation plays a larger role in the diversity of microbiota in each somatic niche. It is therefore no surprise that twin studies have already made prominent contributions to this field.

A study of the gut microbiome in 416 MZ and DZ adult twin pairs found that host genetics plays a role in the frequency of numerous bacterial species, even though bacteria vary widely in heritability (Goodrich et al., 2014). However, a similar study of the dental plaque microbiome from 242 MZ and DZ twin pairs in childhood found that although most variation in the microbiome was determined by environmental factors, highly heritable bacterial species were also present (Gomez et al., 2017). Another recent study investigated MZ twins in Malawi who were discordant for *kwashiorkor*, a malnutrition disorder (Smith et al., 2013). When germ-free mice were given a combination of the Malawian diet and fecal transplant from either affected or healthy twins, *kwashiorkor* developed only in mice that received the transplant from affected twins. This finding implicated microbial imbalance as a major causal factor for this disorder. Using a co-twin control model, Cernada et al. (2016) studied bacterial DNA and gene expression in exfoliated host epithelial cells in the gut from DZ twins discordant for neonatal sepsis. They found that gut microbiomes differed within pairs at birth, and that microbiota in affected twins induced the expression of genes involved in inflammatory and oxidative stress pathways in gut epithelial cells. A similar co-twin model was used to investigate neonatal fecal microbiomes in a pair of same-sex dichorionic twins of unknown zygosity who were discordant for necrotizing enterocolitis, a serious disease that occurs when intestinal tissues become damaged and begin to die (Hourigan et al., 2016). Results showed that the twins had distinctly different microbiomes, and that changes in microbiome complexity preceded clinical symptoms. Large, longitudinal studies of twins with a range of pathologies will doubtless reveal much more about the bacterial, viral and fungal origins of disease.

## Stem Cell Models of Disease

Patient-derived induced pluripotent stem cells (those capable of giving rise to several different cell types) are used with increasing frequency to model disease *in vitro*. In this method, patient cells, typically from skin or blood, are de-differentiated *in vitro* to a state in which they can be re-differentiated into disease-relevant cell types while still retaining genetic and possibly epigenetic signatures of disease. Stages of early development can then be recapitulated in the desired cell type, and genotype–phenotype associations can be quantified. To our knowledge, only three studies have adopted this approach with twins.

In a study of neurons derived from stem cells contributed by MZ twins discordant for Parkinson's disease, the affected twin's neurons showed lower levels of dopamine, increased expression of the monoamine oxidase gene whose product degrades dopamine and impaired neural network activity (Woodard et al., 2014). These findings guided the development of an *in vitro* treatment that could eventually lead to *in vivo* treatment for Parkinson's. Another study of MZ twins discordant for trisomy 21 found epigenetic changes in the promoter regions of genes on other chromosomes involved in the growth of embryonic organs, yielding important insights into the mechanism by which this condition develops (Sailani et al., 2015). Finally, neurons derived from stem cells were used to study MZ twins discordant for Rett syndrome, which affects intellectual and physical development, primarily in females. Findings identified associated this syndrome with defects in the behavior of astrocytes (support cells in the brain) accompanied by changes in the expression of several genes (Andoh-Noda et al., 2015). A surprising result from this work is that epigenetic and genetic differences within MZ pairs were maintained in

lineages derived from stem cells. These and similar results underscore the critical importance of using twin samples in stem cell models of human disease.

### Twin Registries and International Networks

Over the past 36 years, twin registries have become increasingly sophisticated. Today, they play an important role in facilitating twin studies for medical research, especially by recruiting twin participants through population-based and volunteer-based methods. In addition, they increasingly collect and store biological samples along with other twin data, applying complex systems for database management, data handling and confidentiality (Hur & Craig, 2013).

The International Network of Twin Registries was recently formed to foster scientific collaboration and promote twin research globally by expanding the resources of participating twin registries, most notably in the form of a catalogue of twin data and biospecimens available for new twin studies (Buchwald et al., 2014). The Network was created in 2011 as a working group of the International Society of Twin Studies, which was established only eight years before Hrubec and Robinette's review. Network members share expertise in managing twin registries, conducting twin studies and analyzing data.

Working in conjunction with these organizations, the Collaborative project of Development of Anthropometrical measures in Twins (CODATwins) has assembled the largest international dataset ever created for twin analyses. CODATwins unites representatives from more than 40 twin cohorts in 22 countries who share data on more than 400,000 twins, including more than 800,000 measures of height and weight (Silventoinen et al., 2015). Collaborators aim to study the variance components and early-life antecedents of height and BMI, as well as their variation by age, ethnicity and country. To date, researchers with CODATwins have found that DZ twins in childhood are taller and have higher BMI than MZ twins (Jelenkovic et al., 2015), that genetic factors play a major role in the variation of BMI among adolescent populations of different ethnicities exposed to different obesogenic factors (Silventoinen et al., 2016) and that first-born twins are slightly taller and have higher BMI than their second-born co-twins (Yokoyama et al., 2016). Finally, the 'Mothers of Twins Clubs' cited by Hrubec and Robinette (1984) as a key resource to identify twins for research engagement have since evolved into multiple birth associations established across the world, with global representation through the International Council of Multiple Birth Organisations.

### Conclusion

In 1984, Hrubec and Robinette described the application of twin models in medical research. Our review discusses how far this endeavor has progressed over the past three and a half decades by singling out the studies that in our view illustrate the best new work. Twin studies now benefit from an expanded set of statistical models and a concerted global effort to coordinate twin research projects. They have altered the way we think about the etiology of such disorders as epilepsy, autism and schizophrenia. They have also demonstrated that heritability can differ according to age, socioeconomic status and total phenotypic variance. Research involving twins is poised to yield great benefit to medicine, especially by applying findings from twins to non-twin populations in order to illuminate the variety of factors that make us human.

**Acknowledgments.** The authors thank Raymond M. Harris, PhD, for his invaluable reviews of the manuscript. JMC is supported by grants from the Australian National Health and Medical Research Council (grant numbers 1011070 and 1083779), the Financial Markets Foundation For Children (grant number 032-2007) and the Murdoch Children's Research Institute, which is funded by the Victorian Government's Operational Infrastructure Support Program. LCF and JMC are supported by a Centre of Research Excellence Grant (1079102) from the National Health and Medical Research Council of Australia. LCF is supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Ministry of Education, Brazil.

### References

- Allione, A., Marcon, F., Fiorito, G., Guarrera, S., Siniscalchi, E., Zijno, A., ... Matullo, G. (2015). Novel epigenetic changes unveiled by monozygotic twins discordant for smoking habits. *PLoS One*, *10*, e0128265.
- Andoh-Noda, T., Akamatsu, W., Miyake, K., Matsumoto, T., Yamaguchi, R., Sanosaka, T., ... Okano, H. (2015). Differentiation of multipotent neural stem cells derived from Rett syndrome patients is biased toward the astrocytic lineage. *Molecular Brain*, *8*, 31.
- Antoniou, E. E., Derom, C., Thiery, E., Fowler, T., Southwood, T. R., & Zeegers, M. P. (2011). The influence of genetic and environmental factors on the etiology of the human umbilical cord: The East Flanders Prospective Twin Survey. *Biology of Reproduction*, *85*, 137-143.
- Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychological Medicine*, *25*, 63-77.
- Barker, D. J., & Osmond, C. (1988). Low birth weight and hypertension. *BMJ*, *297*, 134-135.
- Bell, J., & Saffery, R. (2012). The value of twins in epigenetic epidemiology. *International Journal of Epidemiology*, *41*, 140-150.
- Berkovic, S. F., Howell, R. A., Hay, D. A., & Hopper, J. L. (1993). Twin birth is not a risk factor for seizures. *Neurology*, *43*, 2515-2519.
- Bonellie, S. R., Currie, D., & Chalmers, J. (2005). Comparison of risk factors for cerebral palsy in twins and singletons. *Developmental Medicine and Child Neurology*, *47*, 587-591.
- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, *3*, 872-882.
- Brodin, P., Jovic, V., Gao, T., Bhattacharya, S., Angel, C. J., Furman, D., ... Davis, M. M. (2015). Variation in the human immune system is largely driven by non-heritable influences. *Cell*, *160*, 37-47.
- Buchwald, D., Kaprio, J., Hopper, J. L., Sung, J., Goldberg, J., Fortier, I., ... Harris, J. R. (2014). International network of twin registries (INTR): Building a platform for international collaboration. *Twin Research and Human Genetics*, *17*, 574-577.
- Buil, A., Brown, A. A., Lappalainen, T., Vinuela, A., Davies, M. N., Zheng, H. F., ... Dermitzakis, E. T. (2015). Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. *Nature Genetics*, *47*, 88-91.
- Carr, A. B., Martin, N. G., & Whitfield, J. B. (1981). Usefulness of the co-twin control design in investigations as exemplified in a study of effects of ascorbic acid on laboratory test results. *Clinical Chemistry*, *27*, 1469-1470.
- Cernada, M., Bauer, C., Serna, E., Collado, M. C., Martinez, G. P., & Vento, M. (2016). Sepsis in preterm infants causes alterations in mucosal gene expression and microbiota profiles compared to non-septic twins. *Scientific Reports*, *6*, 25497.
- Chiarella, J., Tremblay, R. E., Szyf, M., Provencal, N., & Booij, L. (2015). Impact of early environment on children's mental health: Lessons from DNA methylation studies with monozygotic twins. *Twin Research and Human Genetics*, *18*, 623-634.
- Costanzo, A., Nowson, C., Orellana, L., Bolhuis, D., Duesing, K., & Keast, R. (2018). Effect of dietary fat intake and genetics on fat taste sensitivity: A co-twin randomized controlled trial. *American Journal of Clinical Nutrition*, *107*, 683-694.
- Davey, C. G., Lopez-Sola, C., Bui, M., Hopper, J. L., Pantelis, C., Fontenelle, L. F., & Harrison, B. J. (2016). The effects of stress-tension on depression



- and anxiety symptoms: Evidence from a novel twin modelling analysis. *Psychological Medicine*, 46, 3213–3218.
- De Moor, M. H., Boomsma, D. I., Stubbe, J. H., Willemsen, G., & de Geus, E. J. (2008). Testing causality in the association between regular exercise and symptoms of anxiety and depression. *Archives of General Psychiatry*, 65, 897–905.
- Dickinson, J. E., Keil, A. D., & Charles, A. K. (2006). Discordant fetal infection for parvovirus B19 in a dichorionic twin pregnancy. *Twin Research and Human Genetics*, 9, 456–459.
- Dwyer, T., Blizzard, L., Morley, R., & Ponsonby, A. L. (1999). Within pair association between birth weight and blood pressure at age 8 in twins from a cohort study. *BMJ*, 319, 1325–1329.
- Elks, C. E., den Hoed, M., Zhao, J. H., Sharp, S. J., Wareham, N. J., Loos, R. J., & Ong, K. K. (2012). Variability in the heritability of body mass index: A systematic review and meta-regression. *Frontiers in Endocrinology*, 3, 29.
- Gielen, M., Lindsey, P. J., Derom, C., Loos, R. J., Derom, R., Nijhuis, J. G., & Vlietinck, R. (2006). Curves of placental weights of live-born twins. *Twin Research and Human Genetics*, 9, 664–672.
- Gielen, M., Lindsey, P. J., Derom, C., Loos, R. J., Derom, R., Nijhuis, J. G., & Vlietinck, R. (2007). Twin birth weight standards. *Neonatology*, 92, 164–173.
- Gluckman, P. D. (2012). Epigenetics and metabolism in 2011: Epigenetics, the life-course and metabolic disease. *Nature Reviews. Endocrinology*, 8, 74–76.
- Gluckman, P. D., Hanson, M. A., & Buklijas, T. (2010). A conceptual framework for the developmental origins of health and disease. *Journal of Developmental Origins of Health and Disease*, 1, 6–18.
- Gomez, A., Espinosa, J. L., Harkins, D. M., Leong, P., Saffery, R., Bockmann, M., . . . Nelson, K. E. (2017). Host genetic control of the oral microbiome in health and disease. *Cell Host and Microbe*, 22, 269–278.
- Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhan, R., . . . Ley, R. E. (2014). Human genetics shape the gut microbiome. *Cell*, 159, 789–799.
- Gordon, L., Joo, J. E., Powell, J. E., Ollikainen, M., Novakovic, B., Li, X., . . . Saffery, R. (2012). Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence. *Genome Research*, 22, 1395–1406.
- Hall, J. G. (2003). Twinning. *Lancet*, 362, 735–743.
- Halling, C., Malone, F. D., Breathnach, F. M., Stewart, M. C., McAuliffe, F. M., Morrison, J. J., . . . Perinatal Ireland Research Consortium. (2015). Neuro-developmental outcome of a large cohort of growth discordant twins. *European Journal of Pediatrics*, 175, 381–389.
- Hambrick, D. Z., & Tucker-Drob, E. M. (2015). The genetics of music accomplishment: Evidence for gene–environment correlation and interaction. *Psychonomic Bulletin & Review*, 22, 112–120.
- Herskind, A. M., McGue, M., Iachine, I. A., Holm, N., Sorensen, T. I., Harvald, B., & Vaupel, J. W. (1996). Untangling genetic influences on smoking, body mass index and longevity: A multivariate study of 2464 Danish twins followed for 28 years. *Human Genetics*, 98, 467–475.
- Hibbs, A. M., Black, D., Palermo, L., Cnaan, A., Luan, X., Truog, W. E., . . . Ballard, R. A. (2010). Accounting for multiple births in neonatal and perinatal trials: Systematic review and case study. *Journal of Pediatrics*, 156, 202–208.
- Hopper, J. L. (1993). Variance components for statistical genetics: Applications in medical research to characteristics related to human diseases and health. *Statistical Methods in Medical Research*, 2, 199–223.
- Hopper, J. L., & Seeman, E. (1994). The bone density of female twins discordant for tobacco use. *New England Journal of Medicine*, 330, 387–392.
- Hourigan, S. K., Ta, A., Wong, W. S., Clemency, N. C., Provenzano, M. G., Baveja, R., . . . Niederhuber, J. E. (2016). The microbiome in necrotizing enterocolitis: A case report in twins and minireview. *Clinical Therapeutics*, 38, 747–753.
- Hrubec, Z., & Robinette, C. D. (1984). The study of human twins in medical research. *New England Journal of Medicine*, 310, 435–441.
- Hur, Y. M., & Craig, J. M. (2013). Twin registries worldwide: An important resource for scientific research. *Twin Research and Human Genetics*, 16, 1–12.
- Jamieson, D. J., Read, J. S., Kourtis, A. P., Durant, T. M., Lampe, M. A., & Dominguez, K. L. (2007). Cesarean delivery for HIV-infected women: Recommendations and controversies. *American Journal of Obstetrics and Gynecology*, 197, S96–S100.
- Jelenkovic, A., Yokoyama, Y., Sund, R., Honda, C., Bogl, L. H., Aaltonen, S., . . . Silventoinen, K. (2015). Zygosity differences in height and body mass index of twins from infancy to old age: A study of the CODATwins Project. *Twin Research and Human Genetics*, 18, 557–570.
- Kendler, K. S., & Gardner, C. O. (2014). Sex differences in the pathways to major depression: A study of opposite-sex twin pairs. *American Journal of Psychiatry*, 171, 426–435.
- Kendler, K. S., Kessler, R. C., Walters, E. E., MacLean, C., Neale, M. C., Heath, A. C., & Eaves, L. J. (1995). Stressful life events, genetic liability, and onset of an episode of major depression in women. *American Journal of Psychiatry*, 152, 833–842.
- Li, S., Wong, E. M., Bui, M., Nguyen, T. L., Joo, J. E., Stone, J., . . . Hopper, J. L. (2019). Inference about causation between body mass index and DNA methylation in blood from a twin family study. *International Journal of Obesity*, 43, 243–252.
- Little, S. C., Tikhonov, M., & Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell*, 154, 789–800.
- Lorenz, J. M. (2012). Neurodevelopmental outcomes of twins. *Seminars in Perinatology*, 36, 201–212.
- Martin, N., Boomsma, D., & Machin, G. (1997). A twin-pronged attack on complex traits. *Nature Genetics*, 17, 387–392.
- Matamura, M., Tochigi, M., Usami, S., Yonehara, H., Fukushima, M., Nishida, A., . . . Sasaki, T. (2014). Associations between sleep habits and mental health status and suicidality in a longitudinal survey of monozygotic twin adolescents. *Journal of Sleep Research*, 23, 290–294.
- McNamara, H. C., Kane, S. C., Craig, J. M., Short, R. V., & Umstad, M. P. (2016). A review of the mechanisms and evidence for typical and atypical twinning. *American Journal of Obstetrics and Gynecology*, 214, 172–191.
- McRae, A. F., Powell, J. E., Henders, A. K., Bowdler, L., Hemani, G., Shah, S., . . . Montgomery, G. W. (2014). Contribution of genetic variation to trans-generational inheritance of DNA methylation. *Genome Biology*, 15, R73.
- Mikeska, T., & Craig, J. M. (2014). DNA methylation biomarkers: Cancer and beyond. *Genes (Basel)*, 5, 821–864.
- Mohandas, N., Bass-Stringer, S., Maksimovic, J., Crompton, K., Loke, Y. J., Walstab, J., . . . Craig, J. M. (2018). Epigenome-wide analysis in newborn blood spots from monozygotic twins discordant for cerebral palsy reveals consistent regional differences in DNA methylation. *Clinical Epigenetics*, 10, 25.
- Newsome, J., Vaske, J. C., Gehring, K. S., & Boisvert, D. L. (2016). Sex differences in sources of resilience and vulnerability to risk for delinquency. *Journal of Youth and Adolescence*, 45, 730–745.
- Oliveira, V. C., Ferreira, M. L., Refshauge, K. M., Maher, C. G., Griffin, A. R., Hopper, J. L., & Ferreira, P. H. (2015). Risk factors for low back pain: Insights from a novel case-control twin study. *Spine Journal*, 15, 50–57.
- Pietilainen, K. H., Kaprio, J., Rasanen, M., Rissanen, A., & Rose, R. J. (2002). Genetic and environmental influences on the tracking of body size from birth to early adulthood. *Obesity Research*, 10, 875–884.
- Pimentel, J. D., Szymanski, L. J., Samuel, L. P., & Meier, F. A. (2012). Discordant *Streptococcus agalactiae* (Group B streptococcus) gestational infection in monozygotic/diamniotic and dichorionic/diamniotic twins. *Fetal and Pediatric Pathology*, 31, 176–183.
- Pinheiro, M. B., Morosoli, J. J., Ferreira, M. L., Madrid-Valero, J. J., Refshauge, K., Ferreira, P. H., & Ordonana, J. R. (2018). Genetic and environmental contributions to sleep quality and low back pain: A population-based twin study. *Psychosomatic Medicine*, 80, 263–270.
- Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from each other? *Behavioral and Brain Sciences*, 10, 1–16.
- Poulsen, P., Vaag, A. A., Kyvik, K. O., Moller Jensen, D., & Beck-Nielsen, H. (1997). Low birth weight is associated with NIDDM in discordant monozygotic and dizygotic twin pairs. *Diabetologia*, 40, 439–446.
- Protudjer, J. L. P., Lundholm, C., & Almqvist, C. (2015). Asthma and height in twins: A cohort and within-pair analyses study. *Twin Research and Human Genetics*, 18, 142–150.

- Rakyan, V. K., Down, T. A., Balding, D. J., & Beck, S. (2011). Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics*, *12*, 529–541.
- Rappaport, S. M. (2016). Genetic factors are not the major causes of chronic diseases. *PLoS One*, *11*, e0154387.
- Robinson, E. B., Lichtenstein, P., Anckarsater, H., Hapke, F., & Ronald, A. (2013). Examining and interpreting the female protective effect against autistic behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 5258–5262.
- Saccone, G., Saccone, I., & Berghella, V. (2015). Omega-3 long-chain polyunsaturated fatty acids and fish oil supplementation during pregnancy: Which evidence? *Journal of Maternal-Fetal and Neonatal Medicine*, *29*, 2389–2397.
- Sailani, M. R., Santoni, F. A., Letourneau, A., Borel, C., Makrythanasis, P., Hibaoui, Y., . . . Antonarakis, S. E. (2015). DNA-methylation patterns in trisomy 21 using cells from monozygotic twins. *PLoS One*, *10*, e0135555.
- Schiesser, M., Sergi, C., Enders, M., Maul, H., & Schnitzler, P. (2009). Discordant outcomes in a case of parvovirus b19 transmission into both dichorionic twins. *Twin Research and Human Genetics*, *12*, 175–179.
- Schulte, S., Wolffe, J., Schreiner, F., Stoffel-Wagner, B., Peter, M., Bartmann, P., & Gohlke, B. (2016). Birthweight differences in monozygotic twins influence pubertal maturation and near final height. *Journal of Pediatrics*, *170*, 288–294.
- Silventoinen, K., Jelenkovic, A., Sund, R., Honda, C., Aaltonen, S., Yokoyama, Y., . . . Kaprio, J. (2015). The CODATwins project: The cohort description of COLlaborative project of Development of Anthropometrical measures in Twins to study macro-environmental variation in genetic and environmental effects on anthropometric traits. *Twin Research and Human Genetics*, *18*, 348–360.
- Silventoinen, K., Jelenkovic, A., Sund, R., Hur, Y. M., Yokoyama, Y., Honda, C., . . . Kaprio, J. (2016). Genetic and environmental effects on body mass index from infancy to the onset of adulthood: An individual-based pooled analysis of 45 twin cohorts participating in the Collaborative project of Development of Anthropometrical measures in Twins (CODATwins) study. *American Journal of Clinical Nutrition*, *104*, 371–379.
- Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D. I., Cornes, B. K., Davis, C., . . . Kaprio, J. (2003). Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Research*, *6*, 399–408.
- Smith, M. I., Yatsunenko, T., Manary, M. J., Trehan, I., Mkakosya, R., Cheng, J., . . . Gordon, J. I. (2013). Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. *Science*, *339*(6119), 548–554.
- Song, Y. M., Lee, K., & Sung, J. (2015). Changes in weight and cardiovascular disease risk factors in monozygotic twins: The healthy twin study. *Twin Research and Human Genetics*, *18*, 151–157.
- Souren, N. Y., Gerdes, L. A., Lutsik, P., Gasparoni, G., Beltran, E., Salhab, A., . . . Walter, J. (2019). DNA methylation signatures of monozygotic twins clinically discordant for multiple sclerosis. *Nature Communications*, *10*, 2094.
- Spector, T. D. (2015). *The diet myth: Why the secret to health and weight loss is already in your gut*. New York: Overlook Press.
- Stromswold, K. (2006). Why aren't identical twins linguistically identical? Genetic, prenatal and postnatal factors. *Cognition*, *101*, 333–384.
- Turkheimer, E. (2011). Commentary: Variation and causation in the environment and genome. *International Journal of Epidemiology*, *40*, 598–601.
- van Dongen, J., Nivard, M. G., Willemsen, G., Hottenga, J. J., Helmer, Q., Dolan, C. V., . . . Boomsma, D. I. (2016). Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nature Communications*, *7*, 11115.
- van Dongen, J., Slagboom, P. E., Draisma, H. H., Martin, N. G., & Boomsma, D. I. (2012). The continuing value of twin studies in the omics era. *Nature Reviews Genetics*, *13*, 640–653.
- Woodard, C. M., Campos, B. A., Kuo, S. H., Nirenberg, M. J., Nestor, M. W., Zimmer, M., . . . Noggle, S. A. (2014). iPSC-derived dopamine neurons reveal differences between monozygotic twins discordant for Parkinson's disease. *Cell Reports*, *9*, 1173–1182.
- Yelland, L. N., Sullivan, T. R., & Makrides, M. (2015). Accounting for multiple births in randomised trials: A systematic review. *Archives of Disease in Childhood: Fetal & Neonatal*, *100*, F116–F120.
- Yokoyama, Y., Jelenkovic, A., Sund, R., Sung, J., Hopper, J. L., Ooki, S., . . . Silventoinen, K. (2016). Twins' birth-order differences in height and body mass index from birth to old age: A pooled study of 26 twin cohorts participating in the CODATwins Project. *Twin Research and Human Genetics*, *19*, 112–124.
- Yusuf, K., & Kliman, H. J. (2008). The fetus, not the mother, elicits maternal immunologic rejection: Lessons from discordant dizygotic twin placentas. *Journal of Perinatal Medicine*, *36*, 291–296.